DOCUMENT RESUME

ED 347 197 TM 018 703

AUTHOR Medina-Diaz, Maria
TITLE Analysis of a Mathematics Achievement Test via the Linear Logistic Test Model and Quadratic Assignment.
PUB DATE Apr 92
NOTE 39p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests; Algebra; *Cognitive Processes; Cognitive Structures; Comparative Analysis; Difficulty Level; Equations (Mathematics); *Factor Structure; Grade 9; High Schools; *High School Students; Item Response Theory; Mathematical Models; *Mathematics Tests; Matrices; Test Construction; *Test Items
IDENTIFIERS *Linear Logistic Test Model; *Quadratic Assignment; Rasch Model

ABSTRACT

The cognitive structure of an algebra test was defined and validated using the linear test model (LLTM) and quadratic assignment (QA) techniques. The LLTM is an extension of the Rasch model with a linear constraint that describes item difficulty in terms of the cognitive operations required to solve the item. The model permits the specification of the cognitive structure of a test via the weight matrix. In this study, the structure was based on eight production rules that represent the mathematical procedures for solving linear equations with one variable. A 29-item test was constructed and administered to 235 ninth graders in April of 1991. Subjects included 94 low scorers and 141 high scorers. Item response data were analyzed using the LLTM computer program. Estimates of the basic parameters corresponding to each production rule were obtained. A QA confirmatory approach validated the cognitive structure of the test. The proximity matrix represented the examinees' actual use of production rules in solving every pair of items. Cell entries of the structure matrix reflected the common number of production rules required to solve every item pair. The cognitive structure of the test was validated (i.e., examinees solved items using the rules specified in the weight matrix). This supports the basic parameters estimated via the LLTM as quantitative indexes of the difficulty of the cognitive rules in the weight matrix. One figure and seven tables provide study data. There is a 35-item list of references. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
MARIA MEDINA-DIAZ
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED347197

ANALYSIS OF A MATHEMATICS ACHIEVEMENT TEST VIA THE LINEAR LOGISTIC TEST MODEL AND QUADRATIC ASSIGNMENT

María Medina-Díaz

University of Wisconsin-Madison[1]

Paper presented at the Annual Meeting of the
American Educational Research Association
San Francisco, April 1992

[1]This research was part of the doctoral dissertation of the author at the University of Wisconsin-Madison. The author gratefully acknowledges Professor Frank B. Baker for his valuable suggestions during the course of this investigation.

TM018705

BEST COPY AVAILABLE

Abstract

The purposes of this investigation were to define, and to validate the cognitive structure of an algebra test using the linear logistic test model (LLTM) and quadratic assignment (QA) techniques, respectively. The LLTM is an extension of the Rasch model with a linear constraint that describes the difficulty of the item in terms of the cognitive operations required to solve it. The model permits the specification of the cognitive structure of a test via the weight matrix **W**. In this study this structure was based on a set of eight production rules that represent the mathematical procedures employed in solving linear equations with one-variable. A 29-item test was constructed accordingly, and administered to 235 ninth-graders. Item-response data were analyzed using Fischer and Formann's (1972) LLTM computer program. Estimates of the basic parameters corresponding to each of the production rules were obtained.

A QA confirmatory approach was employed to validate the cognitive structure of the test. The proximity matrix **Q** represented the examinees' actual use of the production rules in solving every pair of items. The cell entries of the structure matrix **C** reflected the common number of production rules required to solve every item pair. Following this approach, the cognitive structure of the test was validated (i.e., the examinees solved the items using the set of rules specified in the **W** matrix). Thus, this supports the basic parameters estimated via the LLTM as quantitative indexes of the difficulty of the cognitive rules included in the W matrix.

Analysis of a mathematics achievement test via the linear logistic test model and quadratic assignment

A long-standing challenge has been the development of psychometric methods for assessing individuals' varied and complex abilities, and the integration of such assessment with learning and instruction. More than ever, there is an increasing interest in accounting for the examinees' cognitive processes in solving a given task, rather than only in accounting for the final answer. One way to approach the current testing demands is to design test items that elicit a certain type of cognitive information from the examinees. To do this it is necessary to establish a relationship between the item content and psychometric characteristics, and the cognitive or psychological attributes that it proposes to measure.

Among testing theories, developments in item response theory (IRT) models and techniques are promising for this line of inquiry. The IRT methodologies make it feasible to analyze test items using a psychological theory as a framework to show what attributes of the items lead to differences in the examinees' performance. The researcher, for example, can find differences among items with respect to their underlying psychometric characteristics, and how these features reveal differences in the individual's proficiency in the items. The IRT componential models (e.g., Whitely, 1980; Embretson 1983; Fischer 1973, 1974) make it feasible to accomplish this aim. One of these models, the linear logistic test model, is at the center of the present investigation.

Linear logistic test model

Fischer (1973, 1974) presented one of the first attempts to account for the cognitive qualities underlying test items; that is, the type of knowledge or procedures required to solve the items. Fischer (1973) found that the difficulties of a set of items about differentiation in calculus can be explained through a series of eight basic operations or rules that the examinee must carry out in order to solve a differentiation problem. Based upon a hypothetical or a priori structure of the test, he postulated that the item difficulty could be expressed in terms of the contribution from each separate rule.

Therefore, for every item i a vector of weights is defined as $w_i = (w_1, w_2, ..., w_m)$, where the element in the k-th position denotes the occurrence of the k-th operation in answering the item. The set of all test item vectors can be presented in the form of a test weight matrix **W** with dimensions n x m (m < n), where n denotes the number of items and m denotes the number of operations. Accordingly, the item difficulty parameter is seen as a linear function of the occurrence and difficulty of the necessary operations for solving the task. The resulting componential model is called the "linear logistic test model" (LLTM). Thus, the difficulty of item i is given by

$$\beta_i = \sum_{k=1}^{m} w_{ik}\,\eta_k + c, \quad \text{where} \qquad (1)$$

β_i = item i difficulty parameter (i = 1, 2, ..., n),

w_{ik} = weight (occurrence) with which operations of type k are involved in solving item i,

η_k = difficulty parameter (called "basic parameter") for the cognitive operation k (k= 1, 2,..., m), and

c = normalization constant,

$$c = -\frac{1}{n}\sum_{i}^{n}\sum_{k}^{m} w_{ik}\,\eta_k, \text{ such that } \sum_{i=1}^{n}\beta_i = 0.$$

Mathematically, the LLTM is an extension of the Rasch model with a linear constraint attached (Fischer, 1983). This linear constraint (Equation 1) permits the specification of the contributions of every cognitive operation defined a priori to the difficulty of the item. These cognitive operations are not necessarily single elementary information processes, rather they can be a combination of more elementary processes (Embretson, 1985).

A number of interesting and successful applications have been undertaken using LLTM (e.g., Fischer, 1973, 1978; Spada, 1977; Whitely & Schneider, 1981; Fischer & Formann, 1982; Koponen, 1983; Mitchell, 1983; Hornke & Habon, 1986; Green & Smith, 1987; Embretson & Wetzel, 1987; and Sheehan & Mislevy, 1990). Despite the success of this model, mostly with aptitude tests the application to achievement testing has been limited. Further, the cognitive operations involved in solving the items have been defined in a broad sense. For instance, Fischer (1973) and Spada (1977) defined the hypothesized operations through the topics covered by the items and through the properties or features of the task, respectively, but without a detailed account of the specific cognitive processes underlying those.

In addition, the way in which these operations have been defined and specified for test construction purposes is an area of concern in light of the recent conceptualizations of knowledge and learning developed from cognitive psychology research (e.g.,

Rumelhart and Norman, 1978; Anderson, 1985). The recent evaluation of document literacy skills by Sheehan and Mislevy (1990) using the LLTM constitutes an advance in this regard. They examine these skills through parsing models (i.e., semantic representations) which represent the individual's information-processing mechanisms in reading a given text.

Another critical issue related to the application of the LLTM is the scarcity of approaches to evaluating the validity of the **W** matrix. Statistical procedures such as the likelihood ratio are used to test the overall validity of the LLTM against other models (e.g., the Rasch model). Inferences are made in terms of the adequacy of the model in describing the observed data. However, the LLTM relies on the specification of the **W** matrix but no other statistical procedure has been used to confirm the structure of this matrix; and hence, to support the researcher's initial hypothesis about the cognitive complexity of the test items. If the **W** matrix is false, so will be the estimates of the parameters of the cognitive operations (Fischer, 1977). Thus, the validation of the task structure matrix is an important aspect of the employment of the LLTM for both test analysis and test design.

Quadratic assignment (QA) techniques of confirmatory analysis have been used to verify the theoretical structures in education and psychology (see Hubert & Schultz, 1976; Baker & Hubert, 1977). QA refers to a group of data analysis procedures dealing with object orderings and with the relationship of the proximity matrix to an hypothetical underlying organization of those objects. In the present investigation, QA procedures were employed for determining the presence of the theoretical structure of the test (**W** matrix) in the examinees' actual item solutions that are contained in the proximity matrix **Q**.

The purpose of this study were twofold. First, to define the cognitive structure of an algebra test covering linear equations with one variable employing the LLTM framework. Second, to validate this structure via quadratic assignment procedures.

Methodology

The procedures for carrying out this investigation were divided into two major phases: (a) development and administration of the test, and (b) data analysis.

Development of the test

The first task in this investigation was to construct the items of an algebra test. Five textbook series (Brumfiel, Golden & Heins, 1986; Price, Rath, & Leschensky, 1989; Foster, Rath, & Winters, 1990; Dolciani, Sorgenfrey, & Graham, 1988; Dolciani, Brown, & Cole, 1988) were consulted regarding the skills included in pre-algebra and algebra courses. Based on this review, fifteen items covering linear equations were constructed and administered to ten high school students. The students were asked to think-aloud while they were solving the equations.

Using the skills covered in the algebra textbooks consulted and the protocol data, a set of eight production rules was inferred. These production rules, denoted by P1, P2,..., P8, are shown in Figure 1. It is assumed that each one of these production rules models part of the procedural knowledge required to solve linear equations with one-variable (Mayer, 1980; Lewis, 1981; Sleeman, 1984; Gagné, 1985).

[Insert Figure 1 about here]

A 29 item-test was constructed according to these production rules. Each item requires the use of a particular combination of production rules in order for the equation to be correctly solved. Fourteen of the possible combinations are represented in the 29 items of the test. Table 1 shows the combinations and sequences of production rules needed to solve the items. Each item requires from two to six production rules applied in a distinct sequence. It is assumed that these combinations vary in level of difficulty depending upon the number of production rules included. Five items required two production rules; ten items required three production rules; eight items four production rules; four items five production rules, and two items six production rules.

[Insert Table 1 about here]

Test administration

The test was admistered to 235 students in two high schools during April of 1991. This sample of examinees were ninth graders in introductory algebra classes. Twelve out of 235 examinees answered all the items correctly. The test mean was 23 with standard deviation of 5.38. Ninety-four examinees scored under the test mean, and they were referred to as the low-scoring subgroup of examinees. The mean score of this subgroup was 17 with the standard deviation of 5.22. The high-scoring subgroup included 141 examinees, and their mean score was 27 with standard deviation of 1.48. These statistics show that the overall performance on the test of the two subgroups was different, however, the low-scoring examinees as a group did not perform poorly on the test.

Finally, the test data was tested against the RM assumptions of minimal guessing, and of common discrimination. In order to check the assumption of minimal guessing, the

ICC under the three-parameter model using the low-scoring examinees and the total group were inspected. The ICCs of the test items showed low guessing parameter values using both data sets. Under the IRT three-parameter model, seven items had non-zero guessing parameter estimates. The average guessing parameter estimate of these items was approximately 0.12, and its standard deviation was 0.19. Due to content considerations, these seven items were not eliminated from the test. However, it was concluded that the minimum guessing assumption was met. The three data sets were analyzed under the Rasch model using Fischer and Formann's (1972) LLTM computer program.

The likelihood ratio (LR) test was employed to examine the assumption of common item discrimination parameters of the test items (Andersen, 1973; Gustafsson, 1980). The LR test indicated non-significant differences in the item parameters calculated using the low-scoring and high-scoring examinees (i.e., group with test scores greater than the test mean of 23), and the total sample ($\chi^2_{(28)} = 40.37$, $\alpha > 0.05$). This implies that the assumption of common item discrimination parameters held; therefore, the 29-item test data fit the RM.

Data analysis

The statistical analysis of the test data included (a) the analysis of the examinees responses to the test via the LLTM, and (b) the application of the quadratic assignment procedures to confirm the hypothetical structure of the test.

LLTM test analysis. The application of the LLTM paradigm involved two major parts: (a) defining the weight matrix, and (b) estimating the basic parameters via Fischer and Formann's (1972) LLTM computer program. This program requires the weight matrix **W** and the examinees' dichotomous item responses as input data. Table 2 shows the 29 x 8 weight matrix **W** of the test containing the vectors of the production rules required for solving each of the 29 items of the test. Each position of these vectors indicate the occurrence of a given production rule k in solving an item. The LLTM computer algorithm calculates conditional maximum likelihood (CML) estimates of the basic parameters η_k's as well as of the item difficulty parameters β^*.

[Insert Table 2 about here]

Quadratic assignment. A quadratic assignment (QA) confirmatory analysis (i.e., static case of Hubert & Schultz, 1976) was used as the validation procedure. The purpose of this was to determine whether or not the observed proximity matrix (**Q** matrix) reflects the hypothetical cognitive structure of the test (**C** matrix). This structure was expressed in terms of the eight production rules weights, and represented by the **W** matrix (see Table 2).

In order to perform the QA procedures, the following general steps were carried out: (a) defining the **Q** matrix, (b) defining the **C** matrix, and (c) executing the QA confirmatory procedure using these two matrices.

1. Defining the **Q** matrix

In general, the elements of the **Q** matrix represent a relationship of similarity between every pair of items x,y of the test (i= 1, 2,.., n, $x \neq y$). The **Q** matrix is a n x n

matrix, where both the xth row and the xth column refer to item x, and the entry in row x and column y is denoted by q(x,y). It is also assumed that the diagonal elements are zeros, q(x,x)= 0. In this investigation, the elements of the **Q** matrix indicate the extent to which the same production rules were used by the examinees in solving a pair of test items. The two main stages for computing the elements of the **Q** matrix are described below.

a) Input of the vectors of item responses for each examinee

Since the examinees' actual item responses were considered in the definition of the **Q** matrix their item responses were inspected in order to determine what production rules the examinees actually used. For every examinee, his or her work sheet for each of the 29 items was inspected, every step of the process of solving an equation was examined, and compared with the vector of production rules proposed for solving that particular item (see **W** matrix, Table 2). If an examinee employed a given production rule a value of one was assigned (d = 1) to that rule for that item. A particular production rule was judged as being used in solving the item only when the examinee's written work in the test sheets showed its execution. A value of zero (d = 0) was assigned to a particular production rule when the production rule was not used. Therefore, the selection variables (i.e., the d values) described the examinee's use or non use of each of the eight production rules in solving an item. A computer program written in BASIC was developed to store the vector of production rules that each of the 235 examinees used for solving the 29 items. The program tooks as input for every examinee the item solution in terms of the eight production rules actually used.

b) Calculation of the **Q** matrix cell entries

The **Q** matrix cells included the cross-product of the basic parameters η_k's and the selection variables across examinees. Therefore, a given cell entry q(x,y) of the **Q** matrix corresponding to the items x,y ($x \neq y$) was defined as follows,

$$q(x,y) = \sum_{j=1}^{N} \sum_{k=1}^{m} (d_{kx} d_{ky}) \eta_k \; ; \qquad (2)$$

where d_{kx}, d_{ky} are selection variables (1,0) associated with each k-th production rule, η_k is the estimate of the k-th basic parameter, m denotes the number of basic parameters (k = 1, 2,..,8), and N denotes the number of examinees (j = 1, 2, ..., 235). A computer algorithm written in the TURBO-PASCAL programming language was developed for performing these computations.

2. Defining the structure matrix **C**

The structure matrix **C** was also a 29 X 29 symmetrical matrix. Its cell entries, denoted by c(x,y), contained the sum of the cross product of the weight vectors of the **W** matrix. Thus, the cell elements were computed as follows:

$$c(x,y) = \sum_{k=1}^{m} w_{kx} w_{ky} \; . \qquad (3)$$

where w_{kx} and w_{ky} are the weights of the k-th production rule in the items x and y, respectively. Also, the diagonal elements of the **C** matrix are set to zero, such that $c(x,y)=0$.

The values of the **C** matrix cells ranged from zero to six indicating the number of production rules required in solving every pair of test items. For instance, a value of zero means that the two items are solved using different production rules-- that is, there are no common production rules between those items. A value of six indicates that the solution to the two items involved six common production rules. Therefore, the elements of the **C** matrix are measures of the hypothetical similarity between every pair of test items.

3. QA analysis of the **Q** and **C** matrices

The QA confirmatory approach simply applies a permutation procedure for testing a particular hypothesis--that is, the null hypothesis of random ordering of the rows and columns of the **Q** matrix relative to those of the **C** matrix. The initial **Q** matrix establishes the particular ordering under test and generates the observed Γ value. The index Γ measures the strenght of the association between features of the data contained in the **Q** matrix with the structure matrix **C**. It is expressed as the sum of all cross-products of the corresponding elements in the **Q** and **C** matrices,

$$\Gamma = \sum_{x}^{n} \sum_{y}^{n} [q(x,y)][c(x,y)], \text{ where x and y are two items}$$

($x \neq y$), and q(x,y) and c(x,y) are the cell entries (row x, column y) for these items in the **Q** and **C** matrices, respectively.

A reference distribution of Γ is generated through a permutation procedure based on the n! possible orderings of the rows and columns of the **Q** matrix. The confirmatory analysis requires that the observed value of the index Γ be compared with a randomization distribution for this index. Accordingly, the observed value of Γ and the

critical value derived from the reference distribution are compared in order to test the null hypothesis. If the observed value of Γ falls in the critical region (e.g., $\alpha < 0.05$), the null hypothesis of random allocation is rejected indicating that the **Q** and **C** matrices match on a substantive basis.

Anselin's (1986) MICROQAP computer program was used to perform the QA procedures. Using the **Q** and **C** matrices as the input data, the MICROQAP program executes a series of random permutations (up to 1000) constructing an empirical frequency distribution for the Γ test statistic. In addition to this randomization distribution of Γ (using Monte Carlo methods), MICROQAP computes the Z-unit normal and the Pearson Gamma III distributions as reference distributions. The probability values associated with the reference distributions were used for determining the degree of agreement between the cells of the **C** and **Q** matrices.

A total of four QA analyses were carried out--one using the total sample of examinees (Case 1 as described above), and three using the low- and high-scoring subgroups of examinees.

Case 2. Low-scoring subgroup data (Q matrix) vs. C matrix.

The cells of the **Q** matrix were defined as in Equation 2 but the sum was taken over the 94 examinees of the low-scoring subgroup. The same procedure for calculating the **Q** matrix cell entries as described above was performed. The **C** matrix employed in this case is the the original structure matrix; and its elements were defined as in Equation 3.

Case 3. High-scoring subgroup data (Q matrix) vs. C matrix.

The **Q** matrix cell elements were defined as in Equation 2 but the sum was over 141 examinees (i.e., the number of examinees in the high-scoring subgroup). The original structure matrix **C** was also used in this analysis.

Next, taking as input the **Q** and **C** matrices two separate runs of the computer program MICROQAP were executed--that is, one for each **Q** matrix. The objective of these two QA analyses was to determine whether or not the **Q** matrix based on the performance of the low- and high-scoring subgroups of examinees each reflected the hypothetical structure of the test represented in the matrix **C**.

Case 4. Low-scoring subgroup data (Q matrix) vs. high-scoring subgroup data (C matrix).

The **Q** matrix of the low-scoring subgroup was used. The **C** matrix employed was the **Q** matrix of the high-scoring subgroup. Thus, the **C** matrix contained empirical information rather than a hypothetical structure as was in the previous QA analyses. The aim of using this arrangement was to determine if the responses of the examinees in the low-scoring subgroup (**Q** matrix) present patterns similar to those of the high-scoring examinees (**C** matrix).

Results

Estimates of the basic parameters

The basic parameters η's represent the values associated with each of the eight production rules required in solving the test items. Table 3 shows the estimates of the

eight basic parameters η's and the standard errors of these estimates. These estimates are expressed on the RM original metric and lie on a ratio scale (Fischer, 1977, 1983). The mean of these basic parameter estimates -0.54 with standard deviation of 0.59.

[Insert Table 3 about here]

Seven out of the eight basic parameter estimates were negative, ranging from -0.253 to -1.710. The most negative basic parameter (-1.71) corresponded to the production rule six (P6) that deal with execution of arithmetic operations inside of parentheses. Only the production rule two (P2), which refers to the multiplication of similar elements, obtained a positive estimate of h (0.477), and therefore, the highest difficulty value of the set of basic parameters. The standard error of estimation of the basic parameters ranged from 0.104 to 0.179 indicating an acceptable precision of the estimates.

The numerical values of the basic parameter estimates indicated that the corresponding production rules contributed differently to the item difficulty. Since the solution of a test item involved a particular combination of production rules as shown in the **W** matrix (see Table 1), the difficulty of the items depended on the sum of the basic parameters associated with the production rules (Equation 1). As was expected, a general pattern of increasing difficulty of the items as a function of the number of production rules required to solve them was found. However, there was some overlap between the difficulty estimates associated with the items that include three and four production rules.

Quadratic assignment

Two sets of QA analyses were performed to determine whether the cognitive structure of the test agreed with the patterns of item responses observed in the various data sets. The first set of QA analysis used the item responses of the total sample of examinees (N = 235[1]). The second set of QA analyses were based on the item responses of the low- and high-scoring examinees.

Case 1. QA using the total sample of examinees

The cell entries of the **Q** matrix were defined as in Equation 3, and are shown in Table 4. All cell elements of the **Q** matrix were negative. This is a direct effect of the values of the basic parameters because seven out of eight basic parameters were negative. The size of the **Q** matrix cell elements depended on the contribution of the examinees' responses based on the production rules, and of the basic parameters associated with these production rules. The cell entry values ranged from -28 to -386 with mean value of -202.2 and standard deviation of 74.13.

[Insert Table 4 about here]

The structure matrix **C** representing the hypothetical cognitive structure of the test is shown in Table 5. The mean value of matrix **C** cell entries was 1.91 with standard deviation of 1.13.

[Insert Table 5 about here]

---

[1] The item responses of 235 examinees were used in the QA analyses while data of 223 examinees were employed in the estimation of the basic parameters.

The basal ordering of the **Q** matrix when compared with the **C** matrix yielded an observed Γ value of -352326. The observed value of index Γ exceeded the largest value of the Γ in the sample of 1000 random permutations (α <0.001). The normalized value of Γ was -4.99, using the Z-unit normal distribution as a reference distribution, and a level of 0.05, this value was significant. The correlation between the **Q** and **C** matrices was also significant ($r = -0.58, \alpha < 0.001$). These findings showed that the null hypothesis of random matching of the two matrices was rejected. Thus, the alternative hypothesis of matching based on a substantive relationship between these matrices was accepted.

Case 2. Low-scoring subgroup data (Q matrix) vs. C matrix

The question in this case was whether the empirical item responses of the low-scoring subgroup (**Q** matrix) reflected the hypothetical structure of the test (**C** matrix). The **Q** matrix based on the test performance of the low-scoring subgroup (N = 94) is shown in Table 6. The **C** matrix is the original structure matrix as shown in Table 5.

[Insert Table 6 about here]

The cell entry values of the **Q** matrix ranged from -7.9 to -126.3; the mean was -65.33 and the standard deviation 24.5. The basal ordering of the **Q** matrix when compared with the **C** matrix yielded an observed G value of -113032, and a normalized Γ of -4.66. The correlation between both matrices was -0.53. The probability values associated with the reference distributions of Γ based on the randomization procedure (1000 permutations) and the Z-unit normal distribution were 0.01 and 0.00, respectively. This showed that the null hypothesis of random matching between the cells

of the Q matrix based on the low-scoring subgroup data and the original structure matrix C was rejected at $\alpha = 0.05$.

Case 3. High-scoring subgroup data(Q matrix) vs. C matrix

The QA problem was to determine whether the empirical data based on the item responses of the high-scoring examinees (**Q** matrix) agreed with the hypothetical structure of the test (**C** matrix, Table 5). The **Q** matrix which contains the item responses of the high-scoring examinees (N = 141) is shown in Table 7. The mean of the **Q** matrix elements was -136.9 with standard deviation of 50.57. The cell entries ranged from -18.4 to -287.1.

[Insert Table 7 about here]

The QA results showed an observed Γ index of -239294 and a correlation coefficient between the **Q** and **C** matrices of -0.59. The probability values derived from the reference distributions (i.e., randomization and Z-unit normal) and a significance level of 0.05 showed that the observed value of Γ fell in the critical region of these distributions. Thus, the null hypothesis of random matching between both matrices was rejected.

Case 4. Low-scoring subgroup data (Q matrix) vs. high-scoring subgroup data (C matrix).

The **Q** matrix of the low-scoring examinees and the **C** matrix based on the high-scoring examinees responses are shown in Tables 6 and 7, respectively. These matrices contain empirical item responses of both groups of examinees. The observed association

index Γ was 8211410. This was the largest Γ value obtained from all QA analyses performed. The correlation between both arrangements was 0.94, which was also the highest correlation coefficient found. The probability values associated with the reference distributions of the index Γ showed values close to 0.00, hence were significant at $\alpha = 0.05$. In this case, the null hypothesis of random matching between the **Q** and **C** matrices was also rejected. Thus, the hypothesis of matching between the patterns of item responses of both subgroups of examinees was accepted.

Summary and conclusions

The present investigation focused on defining the cognitive structure of a mathematics test and examining the validity of this structure. This was accomplished by integrating the LLTM and QA procedures in the analysis of the test data. Under the LLTM the cognitive structure of the test was represented by the weight **W** matrix. The elements of this matrix are the one and zero values associated with the occurrence of a given production rule in the item solution. It is assumed that these production rules account for the difficulty of the items as expressed mathematically in Equation 1. The eight basic parameters corresponding to each of these production rules were estimated via CML procedures using Fischer and Formann's (1972) computer program.

In the QA paradigm, the basic parameters were included in the mathematical formulation of the **Q** matrix cells (Equation 2). The **Q** matrix contained information about both the examinees' actual pattern of usage of the production rules in the item solutions, and the basic parameters associated with those rules. The combination of these two elements in the definition of the cell entries of this matrix constituted an empirical model of the examinees' item responses. The QA technique provides a means for

comparing this empirical arrangement with the hypothetical structure of the test contained in the **C** matrix.

The significant agreement between the **Q** and **C** matrices suggests that in fact the patterns of item solutions of the total sample of examinees (**Q** matrix) reflect the hypothetical structure in the **C** matrix. This was also true for the data sets based upon low- and high-scoring examinees' responses. Therefore, the cognitive structure of the test represented by the **C** matrix was reflected in how both the low- and high-scoring examinees solved the test items. This agreement between the responses of the low- and high-scoring subgroups can be influenced by the fact that the overall performance of these groups in the test was relatively high. The mean scores of the low- and high-scoring subgroups was 17 and 26, respectively (i.e., approximately one standard deviation below and above the test mean of 23, respectively).

Since the cells of the **C** matrix were based on the information in the **W** matrix, and the **Q** matrix reflected the structure of the **C** matrix, it can be said that the **Q** matrix also reflected the structure of the **W** matrix. Thus, the cognitive structure of the test was validated through the QA procedures. Consequently, we can conclude that the **W** matrix certainly described the cognitive operations required to solve the test items, and that the basic parameters based on this structure represented these cognitive operations.

In conclusion, the findings of the QA analyses illuminate one of the most critical issues of the application of the LLTM-- the definition and validation of the cognitive structure of the test (**W** matrix). The LLTM permits the researcher to specify a priori the theoretical structure of the test. This structure is taken as the "true" structure because the model does not provide a means to validate it. However, the QA paradigm gives the researcher a statistical tool for confirming the presence of this structure in the examinees responses to the test items. The QA procedures allows one to incorporate

the information about what cognitive operations the examinees used in solving the items, and to compare it with the researcher's theoretical structure of the test. Consequently, the validation of the cognitive structure of the test provides evidence for supporting the basic parameters as a numerical description of the difficulty of the cognitive operations included in the **W** matrix.

References

Anderson, J. R. (1985). *Cognitive psychology and its implications.* 2nd ed. New York: W.H. Freeman.

Andersen, E. B. (1973). A goodness of fit for the Rasch model. *Psychometrika, 38,* 123-140.

Anselin, L. (1986). *MICROQAP, A microcomputer implementation of generalized measures of spatial association.* Santa Barbara, CA: Department of Geography, University of California.

Baker, F. B., & Hubert, L. J. (1977). *Data anlysis: A unified approach.* Madison, WI: Laboratory of Experimental Design, University of Wisconsin.

Brumfiel, V., Golden, N., & Heins, M. (1986). *Pre-Algebra: Skills, problem solving and applications.* Orlando,FL: Harcourt Brace Javonich.

Dolciani, M. P., Sorgenfrey, R.H., & Graham, J. A. (1988). *Pre-Algebra: An accelerated course.* Boston: Houghton Mifflin.

Dolciani, M. P., Brown, R.G., & Cole, M. (1988). *Algebra: Structure and method.* Boston: Houghton Mifflin.

Embretson, S. E. (1985). Multicomponent latent trait models for test design. In S.E. Embretson (Ed.),*Test design: Developments in psychology and psychometrics,* (pp. 195-218). Orlando, Fl: Academic Press.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93,* 179-197.

Embretson, S. E., & Wetzel, D. (1987). Component latent trait models for paragraph comprehension tests. *Applied Psychological Measurement, 11,* 175-193.

Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika, 48,* 3-26.

Fischer, G. H. (1978). Probabilistic test models and their applications. *German Journal of Psychology, 2,* 298-319.

Fischer, G. H. (1977). Linear logistic test models: Theory and application. In W. F. Kempf, & H. Spada (Eds.), *Structural models of thinking and learning* (pp. 203-225). Vienna: Hans Huber.

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests.* (Introduction to the theory of psychological tests). Bern: Huber.

Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica, 37,* 359-374.

Fischer, G. H., & Formann, A. K. (1972). *An algorithm and a FORTRAN program for estimating the item parameters of the linear logistic test model* (Research Bulletin No. 24). Vienna: University of Vienna, Institute of Psychology.

Fischer, G. H., & Formann, A. K. (1982). Some applications of logistic latent trait models with linear constraints on the parameters. *Applied Psychological Measurement, 6*, 397-416.

Foster, A. G., Ruth, J. N., & Winters, L. J. (1990). *Algebra One*. Columbus, OH: Merrill.

Gagné, E. D. (1985). *The cognitive psychology of school learning*. Boston: Little, Brown.

Green, K. E.& Smith, R. M. (1987). A comparison of two methods of decomposing item difficulties. *Journal of Educational Statistics, 12*, 369-381.

Gustafsson, J-E. (1980). Testing and obatining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology, 33*, 205-233.

Hornke, L. F. & Habon, M. W. (1986). Ruled-based item bank construction and evaluation within the linear logistic framework, *Applied Psychological Measurement, 10*, 369-380.

Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical Statistics, 29*, 190-241.

Kieran, C. (1988). Two different approaches among algebra learners. In A.F. Coxford, & Shulte, A. P. (Eds.), *The ideas of Algebra*. Reston, VA: National Council of Teachers of Mathematics.

Koponen, R. (1983). *An item analysis of tests in mathematics applying logistic test models*. Jyväsklä, Finland: University of Jyväsklä.

Lewis, C. (1981). Skill in algebra. In J. R. Anderson (Ed.), *Cognitive skill and their adquisition* (pp. 85-110). Hillsdale, NJ: Lawrence Erlbaum.

Mayer, R. E. (1980). *Different solution procedures for algebra word and equation problems*. Santa Barbara, CA: University of California, Department of Psychology. (ERIC Document Reproduction Service No. ED 205 401)

Mitchell, K. J. (1983). *Cognitive processing determinants of item difficulty on the verbal subtests of the Armed Services Vocational Aptitude Battery*. Arlington, VA: Army Research Institute for the Behavioral and SocialSciences. (ERIC No. ED 260 122)

Price, J., Rath, J. N., & Leschensky, W. (1989). *Pre-Algebra: A problem solving approach*. Columbus, OH: Merrill.

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three models of learning. In J. W. Cotton, & R. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37-53). Hillsdale, NJ: Lawrence Erlbaum.

Sheehan, K., & Mislevy, R. J.(1990). Integrating cognitive and psychometric models to measure document literacy. *Journal of Educational Measurement, 27*, 255-272.

Sleeman, D. (1984). An attempt to understand students' understanding of basic algebra. *Cognitive Science, 8*, 387-412.

Spada, H. (1977). Logistic models of learning and thought. In H. Spada, & Kempf, L. F. (Eds), *Structural Models of Thinking and Learning* (pp. 227-261). Bern: Huber.

Whitely, S.E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika, 45*, 479-494.

Whitely, S. E., & Schneider, L. M. (1981). Information structure for geometric analogies: A test theory approach. *Applied Psychological Measurement, 5*, 383-397.

Figure 1
Operations and productions rules relevant for the solution of linear equations with one-unknown.[1]

Operations

1. Collecting

P1 IF more than one numeral exists on one side of the equation and these numerals are related by addition (or subtraction),
THEN write the total (or difference) of these numbers.

P2 IF there is one numeral and a term on one side of the equation or there are two numbers, and these are related by multiplication,
THEN write the product of those terms or numbers.

P3 If more than one like term exists on one side of the equation, and those terms are related by addition (or subtraction),
THEN write the total (or difference) of those terms.

2. Balancing[2]

P4 IF there is only one term in the equation and the term and a numeral exits on the same side of equation, and they are related by addition or subtraction,
THEN add the additive inverse of the numeral to both sides of the equation.

P5 IF one like term exists on each side of the equation, and such terms are related to numerals by addition (or subtraction),
THEN subtract (or add) the like term on one side of the equation from itself and from the like term on the other side of the equation.

3. Removing parentheses

P6 IF like terms or numerals exist inside the parentheses related by any arithmetic operation,

---

[1] Some of these rules can be composed or linked together by common goal settings and conditions through repeated practice. This can promote a reduction in the number of steps performed by the examinees in solving the equations, and eventually in developing an automatic execution of the algorithm.

[2] It should be noted that the way in which the inverse operations are used in this procedures is different from the way that they are used in the transposing procedure. Though transposing is often considered to be a shortened version of performing the same operation in both sides of an equation, these two procedures may be dissimilar to the students (Kieran, 1988).

THEN execute this operation and write its result.

P7 IF the elements inside the parentheses are not like terms and they are related by addition or substraction, and one term or numberal exits outside of the parentheses,
THEN multiply out the term or numeral by each one of the elements inside the parentheses.

4. Solve for the unknown

P8 IF the goal is to evaluate a variable at one side of the equation, and the variable is related to a numeral by multiplication,
THEN divide by the numeral on both sides of the equation and the result is "variable = quotient".

Table 1
Combinations of production rules of the algebra test

Total production rules	Sequence Production rules	Item Numbers
2	P1, P8	7
	P2, P8	2
	P1, P4	3
	P3, P8	5,15
3	P5, P3, P8	1,10
	P4, P1, P8	4,6,8,13
	P6, P2, P8	9,19,21,24
4	P3, P4, P1, P8	12,16
	P5, P3, P4, P8	14
	P5, P3, P4, P1	17
	P7, P5, P3, P8	20,22,23,25
5	P5, P3, P4, P1, P8	11
	P6, P2, P5, P3, P8	18,27,29
6	P7, P5, P3, P4, P1, P8	26,28

Table 2
Items and the matrix W of the 29 item-test

Items	Production rules 1	2	3	4	5	6	7	8
1) 4x = 2x - 1	0	0	1	0	1	0	0	1
2) 19y =(8)(3)	0	1	0	0	0	0	0	1
3) 12 + c= 4	1	0	0	1	0	0	0	0
4) 7 + 14p = 7	1	0	0	1	0	0	0	1
5) 20a - 9a = 8	0	0	1	0	0	0	0	1
6) 15m + 7 = 6	1	0	0	1	0	0	0	1
7) 5 + 4 = 5a	1	0	0	0	0	0	0	1
8) 2y - 3 = 11	1	0	0	1	0	0	0	1
9) (15 - 9)3x = 1	0	1	0	0	0	1	0	1
10) 7a = -11a + 7	0	0	1	0	1	0	0	1
11) 20 + 10x = 3 - 5x	1	0	1	1	1	0	0	1
12) 15a - 4a + 6 = 9	1	0	1	1	0	0	0	1
13) 5 + 3x - 2 = 7	1	0	0	1	0	0	0	1
14) 3e - 10 = 5e	0	0	1	1	1	0	0	1
15) -1 = 9k - 6k	0	0	1	0	0	0	0	1
16) 17d + 13d - 9= 10	1	0	1	1	0	0	0	1
17) 9n - 3= 8n + 6	1	0	1	1	1	0	0	0
18) 7e(11 - 8)= 7e +3	0	1	1	0	1	1	0	1
19) 6(6y + 3y) = 0	0	1	0	0	0	1	0	1
20) 16x = 7(7x + 3)	0	0	1	0	1	0	1	1
21) m(4 - 1) = 10	0	1	0	0	0	1	0	1
22) 15p = 5(2p - 4)	0	0	1	0	1	0	1	1
23) 7(3c - 5) = 20c	0	0	1	1	1	0	1	0
24) 9 = 7(4c +c)	0	1	0	0	0	1	0	1
25) 3y = (9y -8)3	0	0	1	0	1	0	1	1
26) 9(2 + 5x) = 8x + 2	1	0	1	1	1	0	1	1
27) 9(3p + 4p)= 7 - 7p	0	1	1	0	1	1	0	1
28) (7y - 4)4 =5 +8y	1	0	1	1	1	0	1	1
29) 1 - 3k= 5(17k - 8k)	0	1	1	0	1	1	0	1

W

Table 3
Estimates of the basic parameters (η) for the eight production rules and the standard erros of these estimates

Production	η	Standard errors of estimation
P1	-0.308	0.138
P2	0.477	0.179
P3	-0.253	0.134
P4	-0.300	0.138
P5	-0.516	0.125
P6	-1.710	0.164
P7	-0.754	0.104
P8	-0.921	0.146
Mean	-0.536	0.102
SD	0.590	0.066

Table 4
Q matrix[1] based on the total sample of examinees

Items	1	2	3	4	5	6	7	8	9	10
1	0	194	33	158	241	194	193	196	185	263
2	194	0	32	162	198	200	203	205	124	199
3	33	32	0	95	31	101	88	104	30	34
4	158	162	95	0	158	228	218	235	146	161
5	241	198	31	158	0	196	197	198	187	247
6	194	200	101	228	196	0	263	278	175	200
7	193	203	88	218	197	263	0	267	178	200
8	196	205	104	235	198	278	267	0	180	202
9	185	124	30	146	187	175	178	180	0	190
10	263	199	34	161	247	200	200	202	190	0
11	221	194	96	220	206	258	252	264	179	231
12	214	200	96	224	219	269	262	274	187	224
13	193	201	102	229	195	269	261	279	176	200
14	218	202	54	178	206	210	205	222	180	230
15	238	195	32	158	243	193	193	196	186	244
16	215	200	94	224	219	266	261	271	186	222
17	59	28	65	84	46	94	84	96	30	65
18	223	135	32	140	204	170	170	174	241	230
19	187	134	27	146	191	168	173	176	247	192
20	247	188	33	151	231	189	186	192	217	251
21	208	153	31	152	209	185	187	191	220	211
22	253	198	31	159	242	194	196	202	222	261
23	148	107	31	96	134	116	107	121	138	156
24	222	173	32	160	226	198	199	205	222	227
25	239	190	31	155	229	189	188	195	216	248
26	219	190	89	210	203	247	241	258	213	224
27	247	176	33	157	234	195	195	201	240	254
28	218	190	88	209	205	249	241	257	215	224
29	244	162	30	152	230	188	185	192	227	253

Items	11	12	13	14	15	16	17	18	19	20
1	221	214	193	218	238	215	59	223	187	247
2	194	200	201	202	195	200	28	135	134	188
3	96	96	102	54	32	94	65	32	27	33
4	220	224	229	178	158	224	84	140	146	151
5	206	219	195	206	243	219	46	204	191	231
6	258	269	269	210	193	266	94	170	168	189
7	252	262	261	205	193	261	84	170	173	186
8	264	274	279	222	196	271	96	174	176	192
9	179	187	176	180	186	186	30	241	247	217
10	231	224	200	230	244	222	65	230	192	251
11	0	269	258	232	203	267	119	201	174	218
12	269	0	271	216	220	285	103	191	181	211
13	258	271	0	219	195	266	96	173	171	190
14	232	216	219	0	203	218	67	203	177	216
15	203	220	195	203	0	220	43	204	189	230

[1] All cell entries are negative.

Table 4 Continued

Items	11	12	13	14	15	16	17	18	19	20
16	267	285	266	218	220	0	100	189	183	208
17	119	103	96	67	43	100	0	63	36	59
18	201	191	173	203	204	189	63	0	257	290
19	174	181	171	177	189	183	36	257	0	255
20	218	211	190	216	230	208	59	290	255	0
21	192	201	186	190	209	199	38	252	257	286
22	224	218	198	222	240	215	60	290	261	385
23	146	125	117	155	134	125	57	208	178	292
24	203	215	198	204	226	213	39	255	256	322
25	214	207	190	216	229	208	60	279	248	368
26	268	255	249	224	203	255	110	271	242	353
27	224	218	198	219	234	216	61	310	264	353
28	268	256	253	225	203	255	115	269	245	350
29	215	210	191	213	230	209	67	300	254	325

Items	21	22	23	24	25	26	27	28	29
1	208	253	148	222	239	219	247	218	244
2	153	198	107	173	190	190	176	190	162
3	31	31	31	32	31	89	33	88	30
4	152	159	96	160	155	210	157	209	152
5	209	242	134	226	229	203	234	205	230
6	185	194	116	198	189	247	195	249	188
7	187	196	107	199	188	241	195	241	185
8	191	202	121	205	195	258	201	257	192
9	220	222	138	222	216	213	240	215	227
10	211	261	156	227	248	224	254	224	253
11	192	224	146	203	214	268	224	268	215
12	201	218	125	215	207	255	218	256	210
13	186	198	117	198	190	249	198	253	191
14	190	222	155	204	216	224	219	225	213
15	209	240	134	226	229	203	234	203	230
16	199	215	125	213	208	255	216	255	209
17	38	60	57	39	60	110	61	115	67
18	252	290	208	255	279	271	310	269	300
19	257	261	178	256	248	242	264	245	254
20	286	385	292	322	368	353	353	350	325
21	0	293	197	286	278	271	292	272	276
22	293	0	298	327	381	360	365	360	337
23	197	298	0	232	292	291	268	285	239
24	286	327	232	0	314	306	320	307	296
25	278	381	292	314	0	353	355	351	317
26	271	360	291	306	353	0	342	413	308
27	292	365	268	320	355	342	0	345	362
28	272	360	285	307	351	413	345	0	309
29	276	337	239	296	317	308	362	309	0

Table 5
C matrix

Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	0	1	0	1	2	1	1	1	1	3	3	2	1	3	2	2	2	3	1	3	1	3	2	0	3	3	3	3	3
2	1	0	0	1	1	1	1	1	2	1	1	1	1	1	1	1	0	2	2	1	2	1	0	1	1	1	2	1	2
3	0	0	0	2	0	2	1	2	0	0	2	2	2	1	0	2	2	0	0	0	0	0	1	1	0	2	0	2	0
4	1	1	2	0	1	3	2	3	1	1	3	3	3	2	1	3	2	1	1	1	1	1	1	1	1	3	1	3	1
5	2	1	0	1	0	1	1	1	1	2	2	2	1	2	2	2	1	2	1	2	1	2	1	0	2	2	2	2	2
6	1	1	2	3	1	0	2	3	1	1	3	3	3	2	1	3	2	1	1	1	1	1	1	1	1	3	1	3	1
7	1	1	1	2	1	2	0	2	1	1	2	2	2	1	1	2	1	1	1	1	1	1	0	0	1	2	1	2	1
8	1	1	2	3	1	3	2	0	1	1	3	3	3	2	1	3	2	1	1	1	1	1	1	1	1	3	1	3	1
9	1	2	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	3	3	1	3	1	0	2	1	1	3	1	3
10	3	1	0	1	2	1	1	1	1	0	3	2	1	3	2	2	2	3	1	3	1	3	2	0	3	3	3	3	3
11	3	1	2	3	2	3	2	3	1	3	0	4	3	4	2	4	4	3	1	3	1	3	3	1	3	5	3	5	3
12	2	1	2	3	2	3	2	3	1	2	4	0	3	3	2	4	3	2	1	2	1	2	2	1	2	4	2	4	2
13	1	1	2	3	1	3	2	3	1	1	3	3	0	2	1	3	2	1	1	1	1	1	1	1	1	3	1	3	1
14	3	1	1	2	2	2	1	2	1	3	4	3	2	0	2	3	3	3	1	3	1	3	3	1	3	4	3	4	3
15	2	1	0	1	2	1	1	1	1	2	2	2	1	2	0	2	1	2	1	2	1	2	1	0	2	2	2	2	2
16	2	1	2	3	2	3	2	3	1	2	4	4	3	3	2	0	3	2	1	2	1	2	2	1	2	4	2	4	2
17	2	0	2	2	1	2	1	2	0	2	4	3	2	3	1	3	0	2	0	2	0	2	3	1	2	4	2	4	2
18	3	2	0	1	2	1	1	1	3	3	3	2	1	3	2	2	2	0	3	3	3	3	2	2	3	3	5	3	5
19	1	2	0	1	1	1	1	1	3	1	1	1	1	1	1	1	0	3	0	1	3	1	0	2	1	1	3	1	3
20	3	1	0	1	2	1	1	1	1	3	3	2	1	3	2	2	2	3	1	0	1	3	3	0	4	4	3	4	3
21	1	2	0	1	1	1	1	1	3	1	1	1	1	1	1	1	0	3	3	1	0	1	0	2	1	1	3	1	3
22	3	1	0	1	2	1	1	1	1	3	3	2	1	3	2	2	2	3	1	3	1	0	3	0	4	4	3	4	3
23	2	0	1	1	1	1	0	1	0	2	3	2	1	3	1	2	3	2	0	3	0	3	0	1	3	4	2	4	2
24	0	1	1	1	0	1	0	1	2	0	1	1	1	1	0	1	1	2	2	0	2	0	1	0	0	1	2	1	2
25	3	1	0	1	2	1	1	1	1	3	3	2	1	3	2	2	2	3	1	4	1	4	3	0	0	4	3	4	3
26	3	1	2	3	2	3	2	3	1	3	5	4	3	4	2	4	4	3	1	4	1	4	4	1	4	0	3	6	3
27	3	2	0	1	2	1	1	1	3	3	3	2	1	3	2	2	2	5	3	3	3	3	2	2	3	3	0	3	5
28	3	1	2	3	2	3	2	3	1	3	5	4	3	4	2	4	4	3	1	4	1	4	4	1	4	6	3	0	3
29	3	2	0	1	2	1	1	1	3	3	3	2	1	3	2	2	2	5	3	3	3	3	2	2	3	3	5	3	0

Table 6
Q matrix[1] of the low-scoring examinees

Items	1	2	3	4	5	6	7	8	9	10
1	0.0	66.3	11.3	42.6	81.5	65.9	66.3	65.6	52.3	88.6
2	66.3	0.0	12.0	47.0	72.7	71.8	76.4	75.0	39.0	70.9
3	11.3	12.0	0.0	32.9	11.3	34.0	33.5	36.2	10.4	12.8
4	42.6	47.0	32.9	0.0	44.4	65.0	64.8	68.4	34.3	46.0
5	81.5	72.7	11.3	44.4	0.0	69.6	71.8	70.2	56.7	87.6
6	65.9	71.8	34.0	65.0	69.6	0.0	94.3	96.2	52.7	72.5
7	66.3	76.4	33.5	64.8	71.8	94.3	0.0	97.1	56.5	73.7
8	65.6	75.0	36.2	68.4	70.2	96.2	97.1	0.0	53.4	73.2
9	52.3	39.0	10.4	34.3	56.7	52.7	56.5	53.4	0.0	57.2
10	88.6	70.9	12.8	46.0	87.6	72.5	73.7	73.2	57.2	0.0
11	69.0	67.2	35.1	64.0	69.7	87.7	90.0	90.5	52.3	79.6
12	71.4	71.8	37.1	66.8	77.6	96.0	96.2	98.1	57.7	80.1
13	62.6	70.9	37.7	65.7	67.2	91.5	91.9	97.9	51.6	70.4
14	71.9	75.5	20.9	54.1	73.9	75.0	77.3	82.1	57.1	82.9
15	78.0	68.1	10.4	42.6	84.4	65.0	66.3	65.6	55.1	83.3
16	70.9	71.8	33.4	63.8	75.8	92.9	93.4	94.5	55.2	78.2
17	16.2	10.1	22.6	25.0	16.8	32.2	29.9	32.2	8.9	22.2
18	59.6	41.1	10.4	30.9	56.7	47.5	49.7	48.2	73.5	64.2
19	55.3	42.4	9.2	36.8	60.9	50.6	55.9	54.6	68.5	59.0
20	73.3	60.8	10.1	35.9	72.4	60.4	59.8	59.4	59.4	77.4
21	67.7	54.9	9.8	39.0	69.9	60.5	62.9	63.2	67.3	70.3
22	77.1	69.1	10.1	43.3	80.1	65.6	68.1	69.3	63.9	85.5
23	45.5	38.7	10.6	28.2	46.2	43.4	39.6	44.0	41.1	51.6
24	74.3	63.7	10.4	44.2	79.5	68.3	70.3	71.0	64.5	79.3
25	72.5	64.4	10.7	42.6	73.2	63.5	64.1	65.4	62.4	79.1
26	67.2	64.4	32.8	59.2	67.7	81.5	80.5	87.4	62.1	72.8
27	76.3	61.0	12.2	42.3	77.0	67.1	67.8	68.6	69.9	83.8
28	61.6	61.2	29.7	51.8	65.7	77.8	76.8	79.7	60.5	68.5
29	71.1	54.1	7.9	37.1	72.0	60.3	58.3	60.9	63.2	76.5

Items	11	12	13	14	15	16	17	18	19	20
1	69.0	71.4	62.6	71.9	78.0	70.9	16.2	59.6	55.3	73.3
2	67.2	71.8	70.9	75.5	68.1	71.8	10.1	41.1	42.4	60.8
3	35.1	37.1	37.7	20.9	10.4	33.4	22.6	10.4	9.2	10.1
4	64.0	66.8	65.7	54.1	42.6	63.8	25.0	30.9	36.8	35.9
5	69.7	77.6	67.2	73.9	84.4	75.8	16.8	56.7	60.9	72.4
6	87.7	96.0	91.5	75.0	65.0	92.9	32.2	47.5	50.6	60.4
7	90.0	96.2	91.9	77.3	66.3	93.4	29.9	49.7	55.9	59.8
8	90.5	98.1	97.9	82.1	65.6	94.5	32.2	48.2	54.6	59.4
9	52.3	57.7	51.6	57.1	55.1	55.2	8.9	73.5	68.5	59.4
10	79.6	80.1	70.4	82.9	83.3	78.2	22.2	64.2	59.0	77.4
11	0.0	94.4	86.6	81.7	64.8	90.7	42.4	53.9	51.8	64.0
12	94.4	0.0	97.9	80.1	76.1	102.8	37.6	55.3	55.5	67.4
13	86.6	97.9	0.0	81.0	65.4	91.5	33.6	48.7	51.2	59.7
14	81.7	80.1	81.0	0.0	69.1	79.4	21.3	60.3	58.1	68.0
15	64.8	76.1	65.4	69.1	0.0	75.6	14.4	55.6	56.2	70.6

[1]All cell entries are negative.

Table 6 Continued

Items	11	12	13	14	15	16	17	18	19	20
16	90.7	102.8	91.5	79.4	75.6	0.0	35.0	53.8	54.7	64.1
17	42.4	37.6	33.6	21.3	14.4	35.0	0.0	22.8	13.9	15.9
18	53.9	55.3	48.7	60.3	55.6	53.8	22.8	0.0	78.1	80.1
19	51.8	55.5	51.2	58.1	56.2	54.7	13.9	78.1	0.0	77.7
20	64.0	67.4	59.7	68.0	70.6	64.1	15.9	80.1	77.7	0.0
21	62.1	66.4	59.2	64.1	67.8	64.2	14.4	81.8	89.5	95.4
22	72.4	73.4	66.8	74.9	77.9	69.9	16.4	80.4	83.3	119.4
23	52.4	45.0	43.1	53.0	45.3	43.7	16.2	62.5	58.6	96.8
24	68.0	74.8	66.6	71.6	77.2	73.6	14.6	75.8	84.8	111.1
25	67.5	67.5	62.5	71.0	71.7	67.4	17.5	76.6	74.2	112.2
26	86.6	85.7	81.6	74.2	65.4	82.2	33.5	76.7	74.3	109.5
27	73.0	74.5	67.8	74.4	76.3	71.5	20.0	90.8	82.3	112.5
28	80.8	80.7	79.2	69.6	61.3	78.7	33.1	71.8	73.1	102.7
29	62.2	67.2	60.7	65.3	71.1	65.1	20.4	83.4	76.1	99.9

Items	21	22	23	24	25	26	27	28	29
1	67.7	77.1	45.5	74.3	72.5	67.2	76.3	61.6	71.1
2	54.9	69.1	38.7	63.7	64.4	64.4	61.0	61.2	54.1
3	9.8	10.1	10.6	10.4	10.7	32.8	12.2	29.7	7.9
4	39.0	43.3	28.2	44.2	42.6	59.2	42.3	51.8	37.1
5	69.9	80.1	46.2	79.5	73.2	67.7	77.0	65.7	72.0
6	60.5	65.6	43.4	68.3	63.5	81.5	67.1	77.8	60.3
7	62.9	68.1	39.6	70.3	64.1	80.5	67.8	76.8	58.3
8	63.2	69.3	44.0	71.0	65.4	87.4	68.6	79.7	60.9
9	67.3	63.9	41.1	64.5	62.4	62.1	69.9	60.5	63.2
10	70.3	85.5	51.6	79.3	79.1	72.8	83.8	68.5	76.5
11	62.1	72.4	52.4	68.0	67.5	86.6	73.0	80.8	62.2
12	66.4	73.4	45.0	74.8	67.5	85.7	74.5	80.7	67.2
13	59.2	66.8	43.1	66.6	62.5	81.6	67.8	79.2	60.7
14	64.1	74.9	53.0	71.6	71.0	74.2	74.4	69.6	65.3
15	67.8	77.9	45.3	77.2	71.7	65.4	76.3	61.3	71.1
16	64.2	69.9	43.7	73.6	67.4	82.2	71.5	78.7	65.1
17	14.4	16.4	16.2	14.6	17.5	33.5	20.0	33.1	20.4
18	81.8	80.4	62.5	75.8	76.6	76.7	90.8	71.8	83.4
19	89.5	83.3	58.6	84.8	74.2	74.3	82.3	73.1	76.1
20	95.4	119.4	96.8	111.1	112.2	109.5	112.5	102.7	99.9
21	0.0	101.1	72.9	104.3	91.1	91.4	100.3	89.2	94.2
22	101.1	0.0	99.5	114.5	121.2	115.2	119.7	109.0	107.3
23	72.9	99.5	0.0	86.4	99.3	99.5	90.6	91.5	79.0
24	104.3	114.5	86.4	0.0	108.3	106.3	112.0	103.7	98.5
25	91.1	121.1	99.3	108.3	0.0	110.8	113.8	108.1	95.8
26	91.4	115.2	99.5	106.3	110.8	0.0	114.1	126.0	97.4
27	100.3	119.7	90.6	112.0	113.8	114.1	0.0	112.1	116.8
28	89.2	109.0	91.5	103.7	108.1	126.3	112.1	0.0	94.5
29	94.2	107.3	79.0	98.5	95.8	97.4	116.8	94.5	0.0

Table 7
Q matrix[1] of the high-scoring examinees

Items	1	2	3	4	5	6	7	8	9	10
1	0.0	128.5	22.0	115.7	159.8	128.8	127.4	130.9	132.7	174.8
2	128.5	0.0	20.9	115.7	126.2	128.6	127.4	129.9	85.8	128.5
3	22.0	20.9	0.0	62.8	20.5	67.4	55.4	68.0	20.3	22.0
4	115.7	115.7	62.8	0.0	114.2	163.5	153.6	166.6	112.3	115.7
5	159.8	126.2	20.5	114.2	0.0	127.3	125.2	128.2	130.5	160.1
6	128.8	128.6	67.4	163.5	127.3	0.0	168.7	181.9	122.9	127.9
7	127.4	127.4	55.4	153.6	125.2	168.7	0.0	170.2	122.5	126.8
8	130.9	129.9	68.0	166.6	128.2	181.9	170.2	0.0	126.8	129.7
9	132.7	85.8	20.3	112.3	130.5	122.9	122.5	126.8	0.0	133.4
10	174.8	128.5	22.0	115.7	160.1	127.9	126.8	129.7	133.4	0.0
11	152.3	126.8	61.0	156.2	136.8	170.9	162.8	173.6	127.5	151.5
12	143.5	129.1	59.8	157.2	142.1	173.1	165.9	176.7	129.4	143.8
13	131.2	130.1	64.7	163.6	128.2	178.0	169.6	181.3	125.2	130.3
14	147.0	127.4	33.1	124.4	132.8	135.8	127.7	140.6	123.8	147.6
15	160.1	127.5	22.0	116.3	159.5	128.8	127.4	130.9	131.8	161.3
16	144.4	128.7	61.3	160.5	143.4	173.8	167.7	177.1	130.8	144.0
17	43.1	18.4	42.6	59.2	29.8	62.5	54.8	64.7	21.7	42.9
18	163.5	94.4	22.3	109.6	147.7	122.7	121.2	125.8	168.3	166.5
19	131.8	91.9	18.7	109.3	130.7	118.0	117.5	121.8	179.4	133.0
20	174.4	127.6	22.9	115.7	159.1	129.1	126.4	133.0	158.4	173.7
21	140.9	98.1	21.5	113.5	140.1	125.0	124.3	127.9	153.5	141.2
22	175.9	129.4	21.4	116.0	162.2	129.1	128.0	133.5	158.4	176.3
23	103.0	68.7	21.0	68.7	88.7	73.4	67.8	77.5	97.8	104.8
24	148.4	109.4	21.7	116.6	147.0	129.8	128.9	133.9	157.5	148.5
25	166.6	126.2	21.1	112.9	156.0	126.0	124.3	130.4	154.0	169.0
26	152.2	126.4	57.0	151.0	135.5	165.5	161.0	171.3	151.5	152.0
27	170.7	115.6	21.2	115.4	157.2	128.5	127.4	132.9	171.0	171.0
28	156.6	129.5	59.2	157.4	139.9	171.3	164.4	177.4	155.2	156.4
29	173.6	108.0	22.4	115.4	158.3	128.1	127.4	131.7	164.1	176.7

Items	11	12	13	14	15	16	17	18	19	20
1	152.3	143.5	131.2	147.0	160.1	144.4	43.1	163.5	131.8	174.4
2	126.8	129.1	130.1	127.4	127.5	128.7	18.4	94.4	91.9	127.6
3	61.0	59.8	64.7	33.1	22.0	61.3	42.6	22.3	18.7	22.9
4	156.2	157.2	163.6	124.4	116.3	160.5	59.2	109.6	109.3	115.7
5	136.8	142.1	128.2	132.8	159.5	143.4	29.8	147.7	130.7	159.1
6	170.9	173.1	178.0	135.8	128.8	173.8	62.5	122.7	118.0	129.1
7	162.8	165.9	169.6	127.7	127.4	167.7	54.8	121.2	117.5	126.4
8	173.6	176.7	181.3	140.6	130.9	177.1	64.7	125.8	121.8	133.0
9	127.5	129.4	125.2	123.8	131.8	130.8	21.7	168.3	179.4	158.4
10	151.5	143.8	130.3	147.6	161.3	144.0	42.9	166.5	133.0	173.7
11	0.0	175.3	172.1	151.0	138.5	176.8	77.6	147.2	123.1	154.4
12	175.3	0.0	173.7	136.7	144.8	182.5	65.4	135.8	126.3	144.2
13	172.1	173.7	0.0	138.4	130.4	175.0	62.8	125.1	120.6	130.6
14	151.0	136.7	138.4	0.0	134.4	138.7	46.2	143.3	119.2	148.7

[1] All cell entries are negative.

Table 7 Continued

Items	11	12	13	14	15	16	17	18	19	20
15	138.5	144.8	130.4	134.4	0.0	144.9	28.9	148.9	133.4	160.2
16	176.8	182.5	175.0	138.7	144.9	0.0	65.0	136.0	128.5	144.6
17	77.6	65.4	62.8	46.2	28.9	65.0	0.0	41.0	22.8	44.0
18	147.2	135.8	125.1	143.3	148.9	136.0	41.0	0.0	179.2	210.5
19	123.1	126.3	120.6	119.2	133.4	128.5	22.8	179.2	0.0	177.8
20	154.4	144.2	130.6	148.7	160.2	144.6	44.0	210.5	177.8	0.0
21	130.4	134.8	127.3	126.6	141.6	135.3	24.2	170.2	168.5	190.8
22	152.0	144.8	131.5	147.7	163.0	145.2	43.7	210.5	178.4	266.6
23	94.1	80.0	74.3	102.3	89.5	81.8	41.4	146.4	119.8	195.8
24	135.4	140.2	131.6	133.2	149.1	140.1	25.3	179.3	171.8	211.5
25	146.6	140.4	127.8	145.8	157.3	141.1	43.0	202.7	173.7	256.4
26	181.6	169.8	167.8	150.6	138.5	173.2	76.6	195.2	168.3	244.1
27	151.3	143.5	130.9	145.3	158.6	144.8	41.3	219.3	181.8	241.1
28	187.5	175.3	173.9	155.7	142.3	176.4	82.6	198.0	172.2	248.0
29	153.2	142.9	130.9	147.7	159.2	144.3	47.1	217.2	178.5	225.5

Items	21	22	23	24	25	26	27	28	29
1	140.9	175.9	103.0	148.4	166.6	152.2	170.7	156.6	173.6
2	98.1	129.4	68.7	109.4	126.2	126.4	115.6	129.5	108.0
3	21.5	21.4	21.0	21.7	21.1	57.0	21.2	59.2	22.4
4	113.5	116.0	68.7	116.6	112.9	151.0	115.4	157.4	115.4
5	140.1	162.2	88.7	147.0	156.0	135.5	157.2	139.9	158.3
6	125.0	129.1	73.4	129.8	126.0	165.5	128.5	171.3	128.1
7	124.3	128.0	67.8	128.9	124.3	161.0	127.4	164.4	127.4
8	127.9	133.5	77.5	133.9	130.4	171.3	132.9	177.4	131.7
9	153.5	158.4	97.8	157.5	154.0	151.5	171.0	155.2	164.1
10	141.2	176.3	104.8	148.5	169.0	152.0	171.0	156.4	176.7
11	130.4	152.0	94.1	135.4	146.6	181.6	151.3	187.5	153.2
12	134.8	144.8	80.0	140.2	140.4	169.8	143.5	175.3	142.9
13	127.3	131.5	74.3	131.6	127.8	167.8	130.9	173.9	130.9
14	126.6	147.7	102.3	133.2	145.8	150.6	145.3	155.7	147.7
15	141.6	163.0	89.5	149.1	157.3	138.5	158.6	142.3	159.2
16	135.3	145.2	81.8	140.1	141.1	173.2	144.8	176.4	144.3
17	24.2	43.7	41.4	25.3	43.0	76.6	41.3	82.6	47.1
18	170.2	210.5	146.4	179.3	202.7	195.2	219.3	198.0	217.2
19	168.5	178.4	119.8	171.8	173.7	168.3	181.8	172.2	178.5
20	190.8	266.6	195.8	211.5	256.4	244.1	241.1	248.0	225.5
21	0.0	192.1	124.9	181.6	187.1	179.8	192.4	183.2	182.4
22	192.1	0.0	199.1	212.9	260.4	245.2	245.6	251.5	230.2
23	124.9	199.1	0.0	146.5	193.0	192.1	177.4	194.4	160.9
24	181.6	212.9	146.5	0.0	206.2	200.5	208.4	204.1	197.6
25	187.1	260.4	193.0	206.2	0.0	242.5	241.8	243.5	221.4
26	179.8	245.2	192.1	200.5	242.5	0.0	228.1	287.1	210.8
27	192.4	245.6	177.4	208.4	241.8	228.1	0.0	233.2	245.5
28	183.2	251.5	194.4	204.1	243.5	287.1	233.2	0.0	214.7
29	182.4	230.2	160.9	197.6	221.4	210.8	245.5	214.7	0.0